IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

A Railway Train Number Tracking Method Using a Prediction Approach

QI ZHANG[1,2], ZHIMING YUAN[1,2], LU YAN[3], TAO ZHANG[1,2], YIFENG MIAO[1,2], SHUXIN DING[1,2]

[1]Signal and Communication Research Institute, China Academy of Railway Sciences, Beijing 100081, China
[2]National Research Center of Railway Intelligence Transportation System Engineering Technology, Beijing 100081, China
[3]Science and Information Technology Department, China Academy of Railway Sciences, Beijing 100081, China

Corresponding author: Zhiming Yuan (e-mail: zhimingyuan@hotmail.com).

This work was supported by the National Natural Science Foundation of China under Grant 61790575, Grant U1834211, the Foundation of China Railway Corporation under Grant P2018G008, and the Foundation of China Academy of Railway Sciences Corporation Limited under Grant 2018YJ061.

ABSTRACT Obtaining the accurate real-time train number and location in a railway network is necessary for railway traffic operation control. This paper investigates the implementation and optimization method of train number tracking, by which the train number and location information can be captured. According to the characteristics of the train operation, the train moving trajectory and the train number tracking problem, the mathematical description of the problem and the tracking model based on the railway signaling states and train schedules are proposed. Then, a method using a hidden Markov model prediction is proposed in order to improve the train number tracking correctness. The simulation results are compared with the results obtained under certain restrictions, and the analyses are discussed. The results show that the proposed method can effectively improve the accuracy of train number tracking with better fault-tolerant robustness.

INDEX TERMS Centralized traffic control (CTC), train number tracking, train location prediction, hidden Markov model

I. INTRODUCTION

Centralized traffic control (CTC) systems are one of the most important and widely used systems for train dispatching in Chinese railways, and they help dispatchers make better train operation decisions. Train number tracking (TNT) is the key function of a CTC system, which provides details on the train location with a specified identification (train number) throughout the entire operational area. Moreover, most of the automatic function of a CTC, such as Automatic Route Setting (ARS), Automatic Train Monitoring (ATM), and Automatic Train Arrival/Departure Time Recording (ATA/DTR) all rely on the train number tracking function, which is one of the cores of train control and dispatching. The accuracy of the results calculated by the train number tracking system are critical to the safe and reliable operation of CTC and the dispatchers. Inappropriate decisions may be made based on inaccurate train number and location results, and setting the wrong routes for an incoming train could lead to a disruption state or, even worse, a train collision due to train location detection failure. This is possible in some cases, such as the presence of bad weather, the train weight is insufficient, the track circuit cannot detect the train and the dispatchers are too busy to observe it, etc.

Currently, in traditional railways with fixed block or high-speed railways in a backup mode (if the wireless communication between the train's onboard equipment and the ground devices have broken down), the train number tracking mainly relies on the railway signaling information provided by a computer-based interlocking (CBI) system, a 6502 relay system or a train control center system (TCC), which collects signaling information between the two adjacent station interlocking areas. As the fundamental information for the train number tracking calculation, signaling information transmitted to the CTC is not always consistent with the actual train locations, resulting in inaccurate train number and locations. To improve the reliability, the Wireless Train Number (WTN) system is built on the traditional Chinese railway, through which the train identification and location information from the train's onboard devices can be transmitted to the CTC through the GSM-R wireless channel but with a lower location precision. While this information can be used as auxiliary decision information, it is not very dependable, so the train's real time schedule information is adjusted by the dispatchers from time to time. However, none

of the methods above can guarantee a high accuracy of the CTC train number tracking in a railway environment under the influence of disturbances.

There has been some research on the train number tracking method of a fixed block railway [1]–[4]. For example, [1] used the railway layout topology characteristics with the track circuit state transition to decide the train number and locations; [2] used an expert system that adopted the method of storing the knowledge of the train route setting rules to improve the correctness of the train number tracking between the section parts and the interlocking routes; [3] used finite automaton to model the relationship between the train locations and the states of tracks and points; and [4] used a Bayesian method to calculate the possibility of the train's next step movement. All of the above studies are based on the assumption that the signaling information received by the CTC can perfectly reflect the true location of the train without considering much about the uncertainty of the entire railway environment.

Few works concentrate on the train number tracking reliability problem, even though it is incredibly important. The main reason for this is that, compared with the increasing system complexity and the effort to improve the reliability of train number tracking, it is much easier for dispatchers to manually find and correct any inaccurate train numbers and locations. This is true in railway areas with a low train density where the dispatchers are not busy, while in railway areas with a high train density, train dispatchers need to monitor several types of information simultaneously through CTC screens and monitors, including train locations, speeds, signaling states, block occupation, system devices status, etc. Due to heavy workloads, it is almost impossible to manually guarantee the complete correctness of the train number's indication without any oversights.

To the best of our knowledge, there are no previous prediction approach to deal with train number tracking reliability problem. Hidden Markov model has been widely used as a prediction method and shows great performance. It has been applied in transportation, such as passenger's trip-chains [5], vehicle trajectory [6], vehicle's speed [7], driver's destinations and routes [8], etc. As a result, this paper is the first attempt to apply hidden Markov model in train number tracking reliability problem. This paper proposes a novel approach to handle this uncertainty issue in the train number tracking processes, as well as minimize the risk of inaccurate train numbers. The main contributions of this work are shown in the following:

1) Defined a custom mathematical model of track number tracking based on signaling indication state and real-time train schedule.
2) Proposed a prediction framework with the hidden Markov model and a Bayesian prediction method to improve the reliability of the train number tracking for a fixed block railway.
3) Application of a novel data oscillation and loss improvement algorithm to eliminate the influences of the signaling indication oscillation and improve the robustness and availability of train number tracking.

The remainders of this paper is organized as follows. In Section II, the concept of train number tracking is described and the mathematical model is proposed. In Section III, the prediction model and framework of train number tracking are proposed and explained. In Section IV, the simulation results are presented and analyzed. Finally, Section V concludes the whole paper.

II. CONCEPT AND MODELING OF TNT

Train number tracking uses the train number as the specific train identification to continually indicate the spatial location of a train in a railway in a time series. The reliability of the train number and the train location are key influencing factors of the operation strategy planning and implementation. The train number tracking is the core function of the CTC system.

A. OVERVIEW OF THE CTC SYSTEM

The CTC system is the railway signal technical equipment used to centralize the control of the signal equipment and to command and manage train operations and shunting service [9]. The CTC system is composed of the three following subsystems:

1) Dispatching Center Subsystem

The dispatching center system performs the train operation between the dispatching center and the stations under proper train planning and managing, including whole control zone train location and state monitoring, train operation schedule generating and adjusting, releasing and transmitting dispatch command orders, shunting and maintenance work planning, automatically or manually train route setting, data message processing and exchanging between center dispatcher workstations and stations, train arrival and departure time recording, stations controlled mode adjusting, etc.

2) Station Subsystem

The station subsystem implements the train operation decisions from the dispatching center. Most of the implementations rely on the Station Autonomous Computer (SAC), which is also the key part of the CTC system. The SAC performs the following important functions: Receive and store the train schedules from the dispatching center and train operation instructions from the dispatching center or station train operators on duty, automatically determine conflicting free train routes or receive the train route setting commands from the operators and deliver them to the interlocking system for execution in due time, set train routes based on the stored train schedules and instructions in case of a disconnection with the dispatching center, solve the confliction between the train operation and shunting operation, collect the signaling equipment status information in real time, track the train number, check the signaling conditions based on the station rules and warn about abnormal conditions.

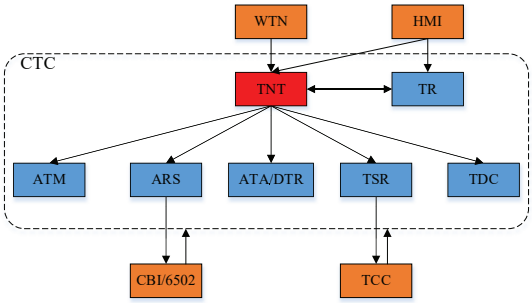


FIGURE 1. Functional structure of the CTC system.

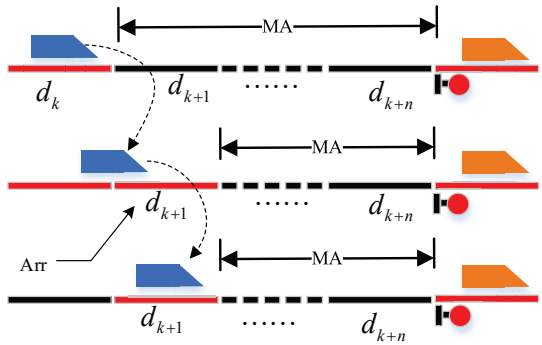


FIGURE 2. Real-time schedule guidance.

3) Network Subsystem

The network subsystem is composed of network communication equipment and transmission channel, which establishes connections and data communications between the dispatching center subsystem and the station subsystem. The network subsystem uses the double-loop and circuitous rings redundancy structure to improve the data transmission reliability.

B. TRAIN NUMBER TRACKING

The train number tracking mainly calculates the train number and location based on the signaling status information received from CBI/6502 and TCC, then transmits the results to the other components of CTC system. Figure 1 refers to the functional structure of the CTC system, and presents the logic relationship between the TNT and other functional parts of the CTC system.

As presented in Figure 1, TNT calculates the train number and locations through the information from CBI/6502, TCC, and WTN. ATM displays the train numbers and locations on the screens. ARS determines the objective routes based on the train numbers and locations and train schedules. ATA/DTR receives the arrival and departure times of a specific train from TNT and records it. Temporary speed restriction (TSR) decides the speed restriction with consideration of the train type (indicated by the train number) and location. The train dispatching command (TDC) formulates the train dispatching command orders based on the information from the TNT, the train regulation adjusts the real-time train schedules based on the exact train number and location, and HMI offers an interface between the operators and TNT to modify the incorrect train numbers.

The train number tracking depends primarily on the persistence signaling information. Combined with auxiliary information, such as real-time schedules, the wireless train number can infer the actual train and location dynamically and accurately. The inference foundation can thus be derived from the following characteristics:

1) Space-time exclusivity

Train $x_j \in X$ has a certain spatial location $d_k \in D$ of absolute exclusive possession at any time point $t_i \in T$, where X is the set of train, D is the set of location and T is the set of time point. In the ideal environment without any signaling equipment failure, external interference or maintenance work, the occupancy state indicated by the signaling information can be considered to be a sufficient and necessary condition for train detection. Let $d_{i,j,k}$ represent the relation of x_j and d_k at t_i, where $d_{i,j,k} = 1$ means x_j on d_k at t_i, otherwise $d_{i,j,k} = 0$. To meet the exclusivity attributes, the following restrictions should be applied:

$$\sum_{j=1} d_{i,j,k} \leq 1 \quad \forall t_i \in T, \forall x_j \in X, \forall d_k \in D. \tag{1}$$

2) Space-time continuity

If train x_j travels from d_k to d_{k+n}, then x_j needs to occupy blocks $d_k, d_{k+1}, d_{k+2}, \ldots, d_{k+n}$ in turn within a certain period of time, as shown in Figure 2. Since the current location of x_j only depends on its last location, the future location only depends on the current location and is independent of the past location, which means the train tracking is in accordance with Markov rules. Let $\tilde{d}_{i,j,k} \in \{0, 1\}$ represent that d_k was either occupied by x_j as of t_i or not. When x_j occupies d_l at t_i, the following restrictions should be applied:

$$\begin{cases} \sum\limits_{u=k}^{l} \tilde{d}_{i,j,u} = l - k + 1 \\ \sum\limits_{u=l+1}^{k+n} \tilde{d}_{i,j,u} = 0 \\ d_{i,j,l} = 1 \\ k < l < k + n. \end{cases} \tag{2}$$

With characteristics 1) and 2), we can establish the mapping relationship between the signaling indications and the train locations, which means that the status changes of the signaling equipment (track, point, block, etc.) that the train travels through reflects the train traveling path and can be used to calculate the train location and determine the train number. Let D represent the signaling equipment set, $d \in D$. Define the binary function $f(d_i, d_j) = \{0, 1\}$, where 1 means d_i and d_j are neighbors, and 0 means otherwise. $v(d_i) = \{0, 1\}$, where 1 means d_i is occupied, and 0 means otherwise. Let $Tr(d_i)$ represent the labeled train number at

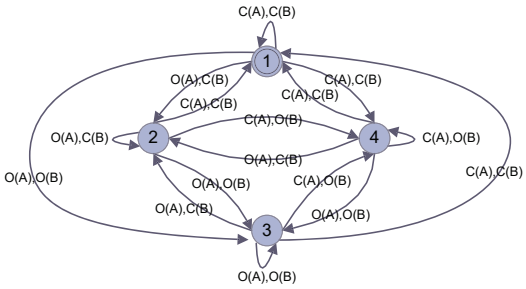


FIGURE 3. Train number tracking state transition.

TABLE 1. Disturbance scenarios.

Disturbances	Signaling states indication
The train is lightweight	Train on but no occupied information, discontinuous occupation
Loss of shunting	Train on but no occupied information, discontinuous occupation
Signaling equipment failure	Wrong signaling information, no train but occupied, train on but no occupied information, etc.
Communication failure	Lost signaling information, bad network condition, etc.

location d_i when d_i is occupied as indicated by the signaling information, then define the train number tracking function as follows:

$$Tr(d_{i+1}) = Tr(d_i) \times f(d_i, d_{i+1}) \times v(d_i) \times v(d_{i+1}). \quad (3)$$

$Tr(d_{i+1}) = 1$ means the train on d_i is the same train on d_i and shares the same train number; otherwise, $Tr(d_{i+1}) = 0$ means the train on d_i has a different train label from the train on d_{i+1}, which should be labeled with other auxiliary information, such as the train schedules we will discuss later or it should be labeled as an unrecognized train if we still cannot obtain sufficient information to determine the train number. Therefore, the train number tracking can be generalized as a discrete state transition problem. Figure 3 refers to the train number tracking state transition.

In Figure 3, "A" and "B" are two neighbor location units in the train traveling path, and the train travel direction is from "A" to "B". State ① represents no train on A and B. State ② represents a train on A but not one on B. State ③ represents a train on A and B. State ④ represents a train on B but not one on A. $O(d_k)$ denotes that d_k is occupied, as indicated by signaling information. $C(d_k)$ denotes that d_k is not unoccupied, as indicated by signaling information.

Signaling information is always mixed with many disturbances when a train is traveling on a railroad in an actual environment, which leads to an uncertain relationship between the signaling information and train location, and sometimes results in the wrong location calculation or a mislabeled train number. Table 1 presents some disturbance scenarios, in which the train location cannot be correctly determined by the signaling information and results in the train being mislabeled or with an abnormal location.

With the following characteristics, the train schedules can be used to improve the robustness and the recovery ability of train number tracking.

3) Train schedules time guidance

Following schedules to dispatch trains is still the basic principle of current railway dispatching operations. At the current time point within a certain range, the real-time schedules have a high accuracy (such as within half an hour; Chinese railways usually adopt a 3-hours real-time schedule. The dispatcher adjusts the time-closer schedules from time to time try to guarantee their correctness). The closer to the time point, the more accurate the train schedules, which means the trains arriving time point and departing time point close to the current time point decided by the schedules have relatively high accuracies. Therefore, the real-time train schedule can be introduced as a way for train number verification. Let $t^a_{j,k}$ and $t^d_{j,k}$ respectively represent the scheduled arriving time and departing time of location d_k for train x_j. Assuming this, x'_j is a scheduled train but we do not know the exact train number labeled by the schedules, and now x'_j occupies track d_k at t_i. If x'_j follows the adjusted real-time schedules, with d_k and t_i as the key features, a unique train can be found in the schedules and should be matched with x_j. To label x_j with the schedules, there are two restrictions that should be met:

a. t_i should be in the scheduled time window $\left[t^a_{j,k}, t^d_{j,k}\right]$, which means that in this period there should be a train staying in the location if the schedules are correct as follows:

$$\left(t^d_{j,k} - t_i\right) \times \left(t_i - t^a_{j,k}\right) \geq 0$$

b. There must be a train x_j that occupies d_k at t_i, which means $d_{i,j,k} = 1$. There should also be a train staying in location d_k, which can be deduced from signaling information.

The train may arrive earlier or departure later, so sometimes the real-time schedules are not accurate. The constants α_1 and α_2 are introduced to obtain better robustness, which respectively represent the train arriving and departing time deviations and should be less than half of the minimum train tracking interval. The time window therefore changes to $\left(t^a_{j,k} - \alpha_1, t^d_{j,k} + \alpha_2\right)$. Let $y_{i,j,k} = 1$, for $\left(t^d_{j,k} + \alpha_1 - t_i\right) \times \left(t_i - t^a_{j,k} + \alpha_2\right) \geq 0$, and 0 otherwise. This leads to the train verification function as follows:

$$Tr(d_k) = \begin{cases} x_j & \text{if } y_{i,j,k} \times d_{i,j} > 0, \forall i, j, k \\ 0 & \text{otherwise} \end{cases}. \quad (4)$$

4) Train schedules sequence guidance

The trains traveling sequences are determined by the schedules. If two trains travel on the same path during a certain

period, the train travel sequence determined by the schedules can be used for train number verification as follows:

$$Tr(d_k) = \begin{cases} x_j & \text{if } f'(x_j, x_{j+1}, t_i) = 1, d_{i,j,k} \times d_{i,j+1,k+n} = 1, \\ & \sum_{j} \sum_{l=k+1}^{k+n-1} d_{i,j,l} = 0, n > 1, \forall i, j, k \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

where $f'(x_j, x_{j+1}, t_i)$ is a function to determine whether or not two trains have the same schedule paths at t_i. equation (5) means there should be no occupied signaling information between d_k and d_{k+n} if x_j and x_{j+1} are following the train according to the schedules at t_i and there should be at least one free block between the two trains. Otherwise, it is difficult to determine if there is one train or two different trains on the two adjacent occupied blocks only based on the schedules.

Although equation (4)-(5) can be used to verify and correct the mislabeled train number caused by an uncertain relationship between the signaling information and the actual train location based on the characteristics of short-term real-time schedules with a high reliability, it is assumed that the schedules will be adjusted in time according to the actual traffic conditions to ensure the train operation is always under control. However, sometimes the schedules may not be updated in time because of a heavy workload, some unexpected events occurring or some other disturbances, such as the train being delayed but the schedules remain the same or the staff having insufficient experience to perform the train regulation and schedule adjustment, etc. This may increase the risk of wrong schedules being made and lead to an uncertainty between the schedule time and the actual train arrival and departure time, which then may result in a mislabeled train number if the schedules are for used train number verification.

III. PREDICTION MODELING AND FRAMEWORK OF THE TNT

A. PREDICTION MODELING OF THE TNT

To improve the results of the train number tracking model based on equation (1)-(5), the uncertainty of the train's real location caused by the randomness of real-time signaling information and schedules, which are the main causes of mislabeled train numbers and locations, should be minimized. Train number tracking can be generalized as a stochastic process of discrete events in a time series and described with a hidden Markov model (HMM). An HMM is a doubly stochastic process with an underlying stochastic process that is not observable (it is hidden). It can only be observed through another set of stochastic processes that produce the sequence of the observed symbols [10].

As presented in Figure 4, the train movement is an underlying stochastic process, which is in accordance with the Markov assumption (future location only depends on the current location of the train, independent of the past location) and can be observed through the signaling state, which is a stochastic process producing the sequence of the observed symbols (the signaling information, etc.). By determining the transition possibilities between the real-train location (hidden states) and the mapping information (observed state), the prediction state (train on or not, same train or not) can be obtained.

Indirectly inferred through the signaling equipment events, the train number tracking scenarios can be identified as the discrete events set in a time series. The relationship between the hidden states and the observed states could be simply captured by the probability matrix. The HMM model of TNT could be defined as follows [10]:

1) Hidden states set

The series of corresponding unobserved train moving trajectories that emit observed movement feature vectors, can be defined as $H = h_0, h_1, \ldots, h_t$, with each hidden state h_t coming from a finite set of N states: $h_t \in S$, $S = \{s_0, s_1, \ldots, s_{N-1}\}$, which can, for instance, be in accordance with the states ① ∼ ④ defined in Figure 3.

2) Observation states set

This refers to the series of observed train movement features, defined as $V = v_0, v_1, \ldots, v_t$, where each observation v_t is a vector consisting of serval dimensions of observed train movement features selected in the HMM model (such as signaling information or real-time schedules).

3) State transition matrix

$$A = \begin{bmatrix} a_{h_1h_1} & a_{h_1h_2} & \cdots & a_{h_1h_n} \\ a_{h_2h_1} & a_{h_2h_2} & & a_{h_2h_n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{h_nh_1} & a_{h_nh_2} & \cdots & a_{h_nh_n} \end{bmatrix}$$

where $a_{h_ih_j} = a_{ij} = P(h_{t+1} = s_j | h_t = s)$, representing the probability that the system goes from state s_i to s_j.

4) Observation probability

Given the initial state at t_0 and all other states until t_i, let $B(d_{i,j,k}) = p(d_{i,j,k}|s_0, \ldots, s_t)$ represent the possibility that train x_j is on location d_k at time t_i; $P(d_{i,j,k}|d_{i-1,j,k-1})$ represents the possibility that x_j is on d_k at t_i given the last state where x_j is on d_{k-1} at t_{i-1} and the observed state changed to a_{k-1} at t_i; $P(z_{i,j,k}|d_{i,j,k})$ represents the possibility that x_j is on d_k but the predicted location is z_k at t_i. If the observed state changes to a_{k-1} at t_i and $z_{i,j,k}$ is not yet estimated, the priori confidence can be expressed as follows:

$$B^-(d_{i,j,k}) = p(d_{i,j,k}|A(z_{i-1,j,k-1}, a_{i-1})), \tag{6}$$

where $A(z_{i-1,j,k-1}, a_{i-1}) = \{z_{0,j,0}, a_0, z_{1,j,1}, a_1, \ldots, z_{i-1,j,i-1}, a_{i-1}\}$. Once $z_{i,j,k}$ is estimated by a_{k-1} with equation (1)-(5), the posterior confidence can be expressed as follows:

$$B^+(d_{i,j,k}) = p(d_{i,j,k}|A(z_{i-1,j,k-1}, a_{i-1}), z_{i,j,k}). \tag{7}$$

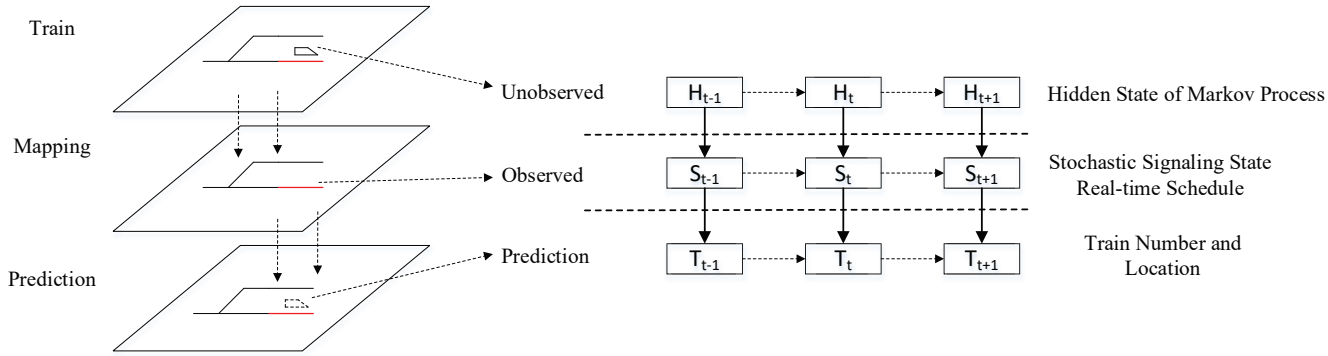


FIGURE 4. HHM of the train number tracking.

According to the Bayes' theorem, the priori confidence can be expressed as follows:

$$B^-(d_{i,j,k}) = \sum_E P(d_{i,j,k}|d_{i-1,j,k-1}, A(z_{i-1,j,k-1}, a_{i-1})) P(d_{k-1}|A(z_{i-1,j,k-1}, a_{i-1}). \tag{8}$$

The past location of train at t_{i-1} is independent of the current observed state changed to a_{i-1}; thus:

$$B^-(d_{i,j,k}) = \sum_E (P(d_{i,j,k}|d_{i-1,j,k-1}, A(z_{i-1,j,k-1}, a_{i-1})) Bel^+(d_{i-1,j,k-1})), \tag{9}$$

and the past location and future location are independent of one another when the current location is known; thus: $P(d_{i,j,k}|d_{i-1,j,k-1}, A(z_{i-1,j,k-1}, a_{i-1}) = P(d_{i,j,k}|d_{i-1,j,k-1}, a_{i-1})$, then:

$$B^-(d_{i,j,k}) = \sum_E P(d_{i,j,k}|d_{i-1,j,k-1}, a_{i-1}) B^+(d_{i-1,j,k-1}). \tag{10}$$

According to the Bayes' theorem, the posterior confidence can be expressed as follows:

$$\begin{aligned} &B^+(d_{i,j,k}) \\ &= \frac{p(z_{i,j,k}|A(z_{i-1,j,k-1},a_{i-1}),d_{i,j,k})p(d_{i,j,k}|A(z_{i-1,j,k-1},a_{i-1}))}{p(z_{i,j,k}|A(z_{i-1,j,k-1},a_{i-1}))} \\ &= \frac{p(z_{i,j,k}|A(z_{i-1,j,k-1},a_{i-1}),d_{i,j,k})B^-(d_{i,j,k})}{p(z_{i,j,k}|A(z_{i-1,j,k-1},a_{i-1}))}. \end{aligned} \tag{11}$$

The predicted location only depends on the current train location, so $p(z_{i,j,k}|A(z_{i-1,j,k-1}, a_{i-1}), d_{i,j,k}) = p(z_{i,j,k}|d_{i,j,k})$ with equations (10) and (13) as follows:

$$\begin{aligned} &B^+(d_{i,j,k}) \\ &= \frac{p(z_{i,j,k}|d_{i,j,k}) \sum_E P(d_{i,j,k}|d_{i-1,j,k-1},a_{i-1}) B^+(d_{i-1,j,k-1})}{p(z_{i,j,k}|A(z_{i-1,j,k-1},a_{i-1}))}, \end{aligned} \tag{12}$$

where the denominator is a normalized constant to ensure the sum of probability is equal to 1.

Therefore, the posterior confidence could be calculated by $p(z_{i,j,k}|d_{i,j,k})$ and $P(d_{i,j,k}|d_{i-1,j,k-1}, a_{i-1})$, which could be obtained through the training of historical traffic data, and the initial confidence could be set as 1.

The HMM should be trained with a series of observation samples first, which could then be used to predict and identify the train location for a given scenario. To obtain more accurate model parameters, large amounts of historical data need to be used as observation samples, which could be collected from the CTC system and should be preprocessed first for the HMM training.

B. ORIGINAL DATA PREPROCESSING

The CTC data to be collected for the prediction model include static data and the dynamic data. The static data may include the station layout, railway line parameters, the original timetable, the train types and the train parameters, while the dynamic data may include the signaling indication, the real-time schedules, the weather conditions and the train arrival and departure time recording (the final timetable).

1) Data oscillation and loss improvement

The original signaling data collected from the real railway environment, which always includes some disturbance, need to be preprocessed first. Generally, one of the most fatal disturbances is the signaling indication oscillation, which is where signaling indications change repeatedly during a short time because of occasional or sudden failure of a transmission channel, signaling equipment, or other reasons, resulting on single point data loss. The signaling indication oscillation and loss leads to a chaotic internal state transition of the TNT, resulting in a mislabeled train number and abnormal train location detection. Taking the continuity of movement, the speed restriction and the exclusivity of train into consideration, it is impossible for a train to travel from one block to the next block during a very short time, 5 seconds for instance, or for a block to be immediately occupied by another train just after one train left. Therefore, the specific signaling indication should not change quickly and repeatedly during a very short time. To eliminate the

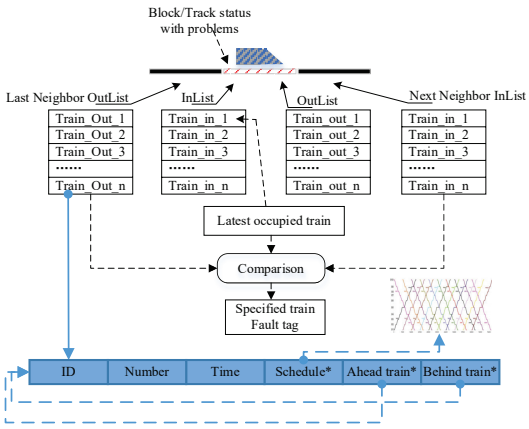


FIGURE 5. Data oscillation and loss improvement algorithm.

influences of the signaling indication oscillation and improve the robustness and availability of TNT, a preprocessed TNT algorithm based on the above analysis is proposed, as shown in Figure 5.

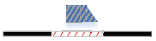
The data structure should be constructed first. Let each train step unit store all passing train information during a specified dynamic moving period of 24 hours. The period is up to the cycle of same train as specified in the timetable in order to avoid storing the same train in the memory. There are two last-in, first-out stacks (InList and OutList) for storing the arrival trains and the departure trains, respectively. The data structure of the stored train information, as presented in Figure 5, includes the train's unique number (ID) in the CTC system, the train number (Number), the arrival time/departure time (Time), the pointer to the specific real-time schedules (Schedule*), the pointer to the train ahead and the train behind that are currently traveling the same route and the same data structure as the stored train information.

Once the collected signaling status information indicates the block/track d_k is occupied but no neighboring train that matches its moving direction can be found, the latest occupied train x_{latest}, which could be obtained from the head of InList, can be resumed as the current occupied train if the following conditions can be satisfied. Otherwise, a fault indication number and an abnormal location warning will emerge. The conditions are as follows:

i. x_{latest} is not in the $d_{k.OutList}$;
ii. x_{latest} and $x_{latest.Train_behind}$ (if not null) are not in $d_{k+i.InList}, 1 \leq i \leq N$, where d_{k+i} matches the moving direction of x_{latest}, and d_{k+N} is the first block/track without an abnormal tag;
iii. $x_{latest.Train_behind}$, if not null, is not in $d_{k-1.OutList}$;
iv. $x_{latest.Train_ahead}$, if not null, is in $d_{k+1.InList}$;
v. $x_{latest.Train_ahead}, x_{latest}, x_{latest.Train_behind}$ are consistent with the schedules.

The amount of resumed train numbers could be counted in the specified period. If the amount exceeds the specified threshold, the block/track could be identified as a fault and a warning will emerge.

TABLE 2. Typical train movement scenes of each step.

No.	Movement	Description
Scene 1		Train on, and indication shows train on
Scene 2		No train but indication shows train on
Scene 3		Train on but indication shows no train
Scene 4		No train but indication shows train on

2) Observation samples acquisition

To obtain the feasible observation samples for scenario training, the filtered signaling data should be carefully classified for serval scenarios. In this paper, based on the relationship between the train movement and the signaling indication, the train movement of each step monitored by the CTC is categorized into four scenes, as shown in Table 2, and the TNT scenario can be developed using simple rules. As most of the railway facilities, such as signaling equipment, infrastructure, trains, timetables, etc., will not change over a long period of time, some specific possible relationships between the train movement and signaling indication in a specific area could be identified from the historical data, which is suitable for HMM training.

Different combinations of any two scenes in Table 2 could characterize the different scenarios of the current movement state of the subject train. If the train's location can be identified, based on which the internal state of the signaling indication memory of TNT could be corrected, the train movement trajectory for the TNT can be generalized as shown in Figure 2.

To acquire more effective observation samples for the HMM train location prediction training, in addition to the signaling indication, other factors, such as real-time schedules, weather conditions, train parameters, and travel times of movement unit should also be considered. The movement unit could be defined as the "signal block", which is the distance between two consecutive signals and can only be occupied by one train at a time, with an entry by another train into it blocked by a protecting red signal. If the train's occupied time of a signal block exceed the limited maximum time, TNT should issue an abnormal train location warning, since the train now could be in an abnormal state caused by train failure, infrastructure failure or operation failure. Figure 6 refers to the method of obtaining the observation samples from historical data collected by the CTC.

With the historical signaling status indication and the railway layout, every signal block status that was changed during the specific history period could be obtained. As in the analysis of section II, the TNT could be designed as an event drive system, and the event window could be determined according to the time of the changed signal block status, which represents the movement step of the train or

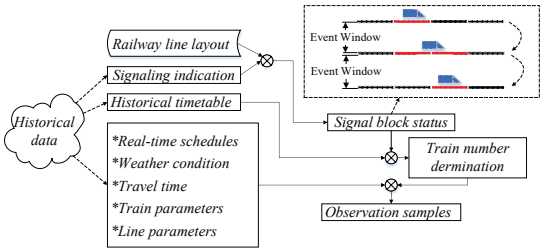


FIGURE 6. Observation sample acquisition.

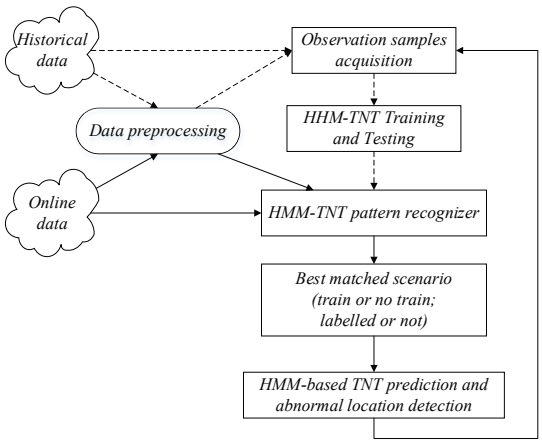


FIGURE 7. Prediction modeling procedure of TNT. Noted that "- - -" indicates offline prediction model training and testing procedure, and "-" indicates online real-time prediction procedure.

the preparation of the train movement. Together with final historical timetable, which correctly records the train movement trajectory with the corresponding train number, the train number of every event window could be determined. Then, with other properties, such as real-time schedules, weather conditions, travel times of the signal block, train parameters and line parameters during the event window period, the observation samples for training could be acquired.

C. PREDICTION FRAMEWORK OF TNT

Based on the analysis above, a prediction framework is proposed here for the TNT development, which consists of two main lines including 1) offline training and testing and 2) online real-time prediction, presented in the flowchart of Figure 7.

1) Offline training and testing

Historical data is first processed to obtain the observation samples (data oscillation elimination and observation samples acquisition procedure as discussed in Section III-B). Then, the observation samples are divided into a training set (75% of all samples) and a testing set (25% of all samples), which are fed into the training of the HMM-TNT pattern recognition model (the detailed algorithm is explained in Section III-A), which constitutes a library of TNT models that would serve as the basis for online prediction.

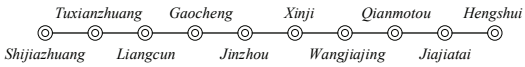


FIGURE 8. Part of the line of the Shijiazhuang-Dezhou Railway for simulation.

The performance of the developed TNT prediction model is finally assessed using the testing samples.

2) Online real-time prediction

Once the HMM model is properly trained offline, online real-time data of the CTC collected after the data processing could be fed to the model to decide the best matched scenario and to predict the train location. Then, the recorded online data could also be added to the offline database later to improve the HMM prediction model.

IV. SIMULATION TEST AND ANALYSIS

To verify the effectiveness and the improvement of the proposed TNT prediction model and framework, simulation experiments are designed with 10 stations (all have at least four arrival and departure tracks, among which the Shijiazhuang and Hengshui stations have more complex station layouts, shown in Figure 8) from the historical CTC data collected from the Shijiazhuang-Dezhou Railway in China. There were 12 passenger trains travelling from Shijiazhuang Station to Hengshui Station, and 16 from Hengshui Station to Shijiazhuang Station every day recorded in historical CTC data with a total of 3120 train steps (1392 steps of Shijiazhuang - Hengshui and 1728 Steps of Hengshui - Shijianzuang). Among which, 15 days of historical data is selected for offline training and 5 days for testing, and the timetable remains the same during these 20 days.

To obtain the performance comparison and verify the improvement of the proposed prediction model and framework, the basic model and the prediction model are programed with the C++ language and respectively tested in the same CTC simulation environment, which has been applied in practice in Chinese railways. Additionally, part of the testing data (such as the signaling equipment status indication) is intentionally modified, such as the fault injection, to make every TNT scenario possible in the simulation environment and under the control. The testing data can then be modified for the details of the data structure that are available. Table 3 presents one group of TNT scenarios for training and testing.

During the training phase, with different recorded parameters from the training samples, the final TNT feature variables selected for training the HMM include the signaling, real-time schedules, traveling times, line parameters and the weather conditions. Table IV presents the training results with the TNT correct rate.

The correct rate R_{TNT} can be calculated as follows:

$$R_{TNT} = 1 - \frac{n_w}{n_{total}}. \tag{13}$$

where n_{total} is the total mislabeled train number if only the signaling TNT algorithm and model are used, and n_w is the

TABLE 3. Testing and training scenarios.

No.	Scenario
S_1	No disturbance.
S_2	Fault zones are created by non-neighbor blocks (each open-track section has one fault zone and each station has two fault zones (inbound and outbound) created by a platform) and no train delayed.
S_3	Fault zones are created by non-neighbor blocks (each open-track section has one fault zone and each station has two fault zones (inbound and outbound) created by a platform) and some trains delayed, but no train is delayed more than the deviation compensations.
S_4	Fault zones are created by non-neighbor blocks (each open-track section has one fault zone and each station has two fault zones (inbound and outbound) created by a platform) and some trains are delayed more than the deviation compensations.
S_5	Fault zones are created by two neighbor blocks and some trains are delayed more than the deviation compensations.
S_6	Fault zones are created by more than two neighbor blocks and some trains are delayed more than the deviation compensations.

TABLE 4. Training results.

No.	Characteristic				
	Signaling	Schedule	Travel time	Line Parameter	Weather
S_1	100%	100%	100%	100%	100%
S_2	79%	100%	100%	100%	100%
S_3	78%	100%	100%	100%	100%
S_4	79%	61%	91%	91%	91%
S_5	47%	43%	78%	87%	87%
S_6	0.9%	5%	15%	21%	21%

Note: From the left to the right of Table IV, each column means one new feature was added for the HMM training, and the raw data presents the correct rate of the TNT tested in the specific scenario.

mislabeled train number during the test, which is counted by the number of changed times of the same mislabeled train during the TNT procedure. The tracked train number changes once, either from the correct to the incorrect or from the incorrect to the incorrect and is then counted as one mislabeled. During the period of mislabeled, it will not be counted if the train number remains the same.

With the simple disturbance scenarios (as S_1, S_2 and S_3), the HMM performs well with few features, and even adds other features that cannot have a better performance. However, as the disturbance scenarios become more complicated (as S_4, S_5 and S_6), the additional features become more important for obtaining a better improvement. It seems that the feature of weather does not cause any improvements from Table IV, because we do not have enough training samples with specific weather conditions. In fact, some extreme weather could cause a signaling indication and train location inconsistency (like heavy rain may cause a loss of shunting). Therefore, we choose to keep the weather feature in the HMM training model.

TABLE 5. Testing results.

No.	Algorithm		
	Signaling	Signaling & Schedule	HMM
S_1	100%	100%	100%
S_2	0%	100%	100%
S_3	0%	75%	100%
S_4	0%	4.3%	92%
S_5	0%	5.1%	87%
S_6	0%	4.9%	23%

During the comparison phase, one day's historical CTC data with a total of 3120 train steps (1392 steps of Shijiazhuang - Hengshui and 1728 Steps of Hengshui - Shijianzuang) is chosen and modified into different versions with different disturbance scenarios and their combinations according to the train number tracking analysis results based on the existing CTC system. Table 5 refers to one of the test results.

Table 5 shows the TNT correct rate with different algorithms. In the case of non-disturbance, the basic signaling model, signaling and schedule model, and the HMM model could all track the train number correctly without any mislabeling. This is because there is a perfect mapping relationship between the signaling status and the actual train location under such conditions. Additionally, the train traveling trajectory is completely consistent with the real-time schedules.

Then, 5 non-neighbor blocks and 5 non-neighbor tracks with a platform were set as the fault zone in turn to generate the disturbance, and the movements of the trains were still consistent with the real-time schedules. The basic signaling model mislabeled a train number whenever a train stepped into these locations, since it is difficult to correctly determine whether or not there is a train on the fault zone only based on the signaling constrains. The correspondence is thus destroyed by the fault zones.

The mislabeled number was reduced by approximately 75% by using the signaling and schedules model under the condition of the 10 fault zones. This is due to the guidance of the real-time schedules, which could be used to determine the train number when the tracks are occupied by the train. These tracks can be considered as the train number check point, which should be carefully selected in practice since it will increase the mislabeled risk if the check point cannot properly match the schedules. Usually, tracking for the train arrival, departure and reversal can be considered as the check point, because a more precise period of track occupancy can be obtained from the real-time schedules. Once the real-time schedules fail to synchronize with the actual train trajectory (train delayed or earlier) and the deviation time exceeds the specific scheduled time window, the schedule guidance will fail, and the train number will not be obtained from schedules or the obtained incorrect train number, which will result in a significant increase in the number of mislabels, as presented

in Table 5.

The deviation compensations (α_1, α_2) of the scheduled time window are important factors for the train number identification. The simulation test shows that the mislabeled train number will increase significantly when the value of the deviation compensations is greater than the headway of the neighbor trains, which will lead to the train number being mislabeled because more than one train will be found in the same scheduled time window. However, if the value of the deviation compensations is too small, there will be no train found in the scheduled time window once the train is earlier or delayed. The determination of the value should be calculated with a distribution of the train delay. In practice, half of the headway is usually a better choice.

There is no mislabeled train number when the HMM model is applied in the simulation environment of the non-neighbor fault zone (none of the mismatched signal blocks are neighbors), which means the model has a good performance in a single point of failure scenario. We then tested the HMM model in the condition of several neighbor blocks being set as the fault zone. As Table 5 shows, the HMM model could label the train correctly in the condition of no alternative successor routes. However, the mislabeled number increased when there were alternative successor routes right next to the fault zone with several neighbor blocks or the number of neighbor blocks in one fault zone increased. In fact, under the condition of serious disturbances, the TNT models listed above will lead to a high risk of the mislabeled train number. The probability relationship of the actual train location and other influencing factors then become extremely complex. Based on the principle of a fail-safe, labeling a fault indication number for the train and providing an abnormal location warning are the better ways to guarantee the safety of the train operation.

V. CONCLUSION

This paper focuses on the train number tracking problem and presents a basic mathematical model, HMM prediction model and a new framework on the basis of the CTC system.

The proposed mathematical model first makes use of the assumption that the train is traveling in an ideal environment where the signaling status information can perfectly reflect the actual train location.

Then, the uncertainty of the mapping relationship between the signaling status information and the train's actual location in the real environment is studied and the real-time schedule is used to verify and correct the train number based on its better accuracy in a specific period close to the current time. To improve the performance of the basic model, an HMM prediction model is proposed based on the theory of HMM, where the observed signaling status information and features could be viewed as the external "performance" of train movement at a specific time, which in essence reflects the internal "hidden states" of the system.

Moreover, a data oscillation and loss improvement algorithm and a new prediction framework are proposed. Finally, simulations of typical train number tracking scenarios are designed and conducted based on the existing CTC system. Promising results are obtained in terms of the train number prediction using the proposed framework based on the simulations.

REFERENCES

[1] D. L. Rush. "Vehicle train tracking apparatus and method," *U.S. Patent* No. 4,361,301. 30 Nov. 1982.
[2] J. Wang, "Railway traffic dispatching control simulation system," *China Railway Sci.*, vol. 28, no. 5, pp. 110–116, May. 2007.
[3] Z. Yuan, C. Chi, J. Guo, "Fault tolerant model of train tracing system for fixed block pattern," *Railway Comput. Appl.*, vol. 22, no. 8, pp. 50–52, Aug. 2013.
[4] Z. Yuan, Q. Zhang, K. Huang, and F. Chen, "A Bayesian Method for Train Number Tracking in Railway Network," *J. Transp. Syst. Eng. Inf. Technol.*, vol. 16, no. 5, pp. 85–89, Oct. 2016.
[5] G. Han, K. Sohn. "Activity imputation for trip-chains elicited from smart-card data using a continuous hidden Markov model," *Transport. Res. B*, 2016, 83: 121-135.
[6] N, Ye, Y. Zhang, R. Wang, et al. "Vehicle trajectory prediction based on Hidden Markov Model," *KSII Trans. Internet Inf.*, 2016, 10(7): 3150: 3170.
[7] B. Jiang, Y. Fei. "Traffic and vehicle speed prediction with neural network and hidden markov model in vehicular networks," *2015 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2015: 1082-1087.
[8] Y. Lassoued, J. Monteil, Y. Gu, et al. "A hidden Markov model for route and destination prediction," *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, 2017: 1-6.
[9] "The Distributed Autonomic Centralized Traffic Control System (CTC) Technology Conditions (Interim)," Science and Technology of Transportation, China Railway Publishing House. No. 15., 2004.
[10] L. Rabiner, and B. Juang, "An introduction to hidden Markov models," *IEEE ASSP magazine*, vol. 3, no. 1, pp. 4–16, Jan. 1986.
[11] S. Eddy, "Hidden Markov Models," *Curr. Opin. Struct. Biol.*, vol. 6, no. 3, pp. 361–365, Jun. 1996.
[12] H. Sun, Z. Hou, and D. Li. "Coordinated iterative learning control schemes for train trajectory tracking with overspeed protection," *IEEE Trans. Autom. Sci. Eng.*, vol. 10, no. 2, pp. 323–333, Apr. 2013.
[13] P. James, et al. "On modelling and verifying railway interlockings: Tracking train lengths," *Sci. Comput. Program.*, vol. 15, no. 96, pp. 315–336, Dec. 2014.

QI ZHANG received the B.S. degrees in Electronic Engineering from Beijing University of Aeronautics and Astronautics, Beijing, China, in 1991, the M.S degrees in Computer Application from China Academy of Railway Sciences, Beijing, in 1993, and the Ph.D. degree in Traffic Information Engineering and Control from China Academy of Railway Sciences in 1998.

He is currently a chief researcher of China Academy of Railway Sciences, and a leader in railway technical expertise. His research interests include railway signal and communication, automatic train operation, train operation control, intelligent dispatching and cooperative control of multiple trains. He has directed many national scientific projects in China and has a lot of achievements. One of his research achievements, the Centralized Traffic Control (CTC) system, has become one of most important technical equipment in China railway.

ZHIMING YUAN received the Ph.D. degree from China Academy of Railway Sciences, Beijing, China, in 2016. He is a Professor of Communication & Signal Research Institute, China Academy of Railway Sciences. His research interests include railway traffic modeling and simulation, railway operations research, and railway signal control.

LU YAN received the B.S. degrees in Automation Science and Electronic Engineering from Beihang University, Beijing, China, in 2005, the M.S degrees in Measuring and Testing Technologies and Instruments from Beihang University, Beijing, in 2008, and in the Ph.D. program of Traffic Information Engineering and Control in China Academy of Railway Sciences.

She is currently an associate researcher of China Academy of Railway Sciences, and a manager in Science and Information Technology Department. Her research interests include railway signal and communication, integrated dispatching, high speed railway earthquake early warning, disaster warning, etc. She has participated many national scientific projects in China and has rich experience in system development.

TAO ZHANG received the Ph.D. degree from China Academy of Railway Sciences, Beijing, China, in 2015. He is an associate Researcher at Researcher of Signal & Communication Research Institute, China Academy of Railway Sciences Corporation Limited. His research interests include railway signal and communication, railway traffic modeling and simulation, automatic train operation, and railway intelligent control and dispatching.

YIFENG MIAO received the B.S. degrees in Computer Applications from Lanzhou Jiaotong University, Lanzhou, China, in 2002, the M.S degrees in Engineering Electronics and Communication Engineering from Beijing Jiaotong University, Beijing, China, in 2011, and the Ph.D. degree in Traffic Information Engineering and Control from China Academy of Railway Sciences in 2015.

He is currently a researcher of Communication Signal Research Institute of China Academy of Railway Sciences Corporation Limited. His research interests include railway communication signal, intelligent scheduling and cloud computing technology.

SHUXIN DING received the B.E. degree in automation and the Ph.D. degree in control science and engineering, both from the Beijing Institute of Technology, Beijing, China, in 2012 and 2019, respectively. From 2016 to 2017, he was a visiting scholar of Industrial and Systems Engineering at the University of Florida, Gainesville, FL, USA.

He is currently a researcher of Communication & Signal Research Institute of China Academy of Railway Sciences. His current research interests include facility location, evolutionary computation, multiobjective optimization, optimization under uncertainty, and railway scheduling.